Math 343 / 643
Spring 2024 (3 credits)
Course Syllabus

Adam Kapelner, PhD.

Queens College, City University of New York

document last updated Tuesday 23rd January, 2024 10:21am

Course Homepage	`https://github.com/kapelner/QC_Math_343_Spring_2024`
Slack Homepage	`https://QCMath343Spring2024.slack.com/`
Contact	@kapelner on slack in a public channel
Office	604 Kiely Hall
Lecture Time and Loc	Tues and Thurs 5:20 – 6:35PM / KY 258
Instructor Office Hours and Loc	see course homepage
TA / Office Hours	see course homepage

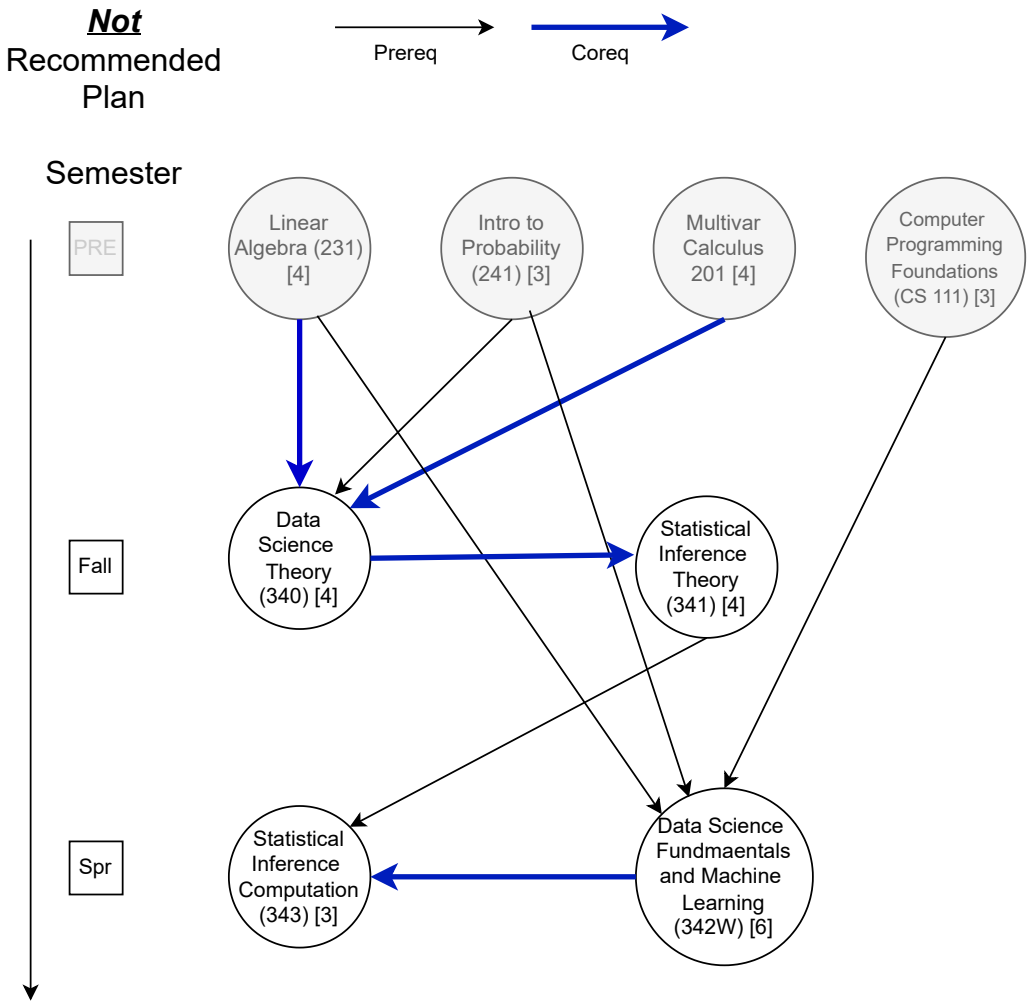
Course Overview

MATH 343 / 643. Computational Statistics for Data Science. 3 hr.; 3 cr. Prereq.: MATH 341 or 641. Coreq.: MATH 342W or 642. Mixture models, EM algorithm, Metropolis-within-Gibbs sampling, permutation tests, the bootstrap, the Kaplan-Meier estimator, the Cox model, T and F tests for the linear model, Gauss-Markov theorem, Bayesian linear regression: Ridge and Lasso. Causality and the randomized experiment, randomization tests. Focus on computation. Special topics. Students cannot receive credit for both: MATH 343 and 643. Fall, Spring

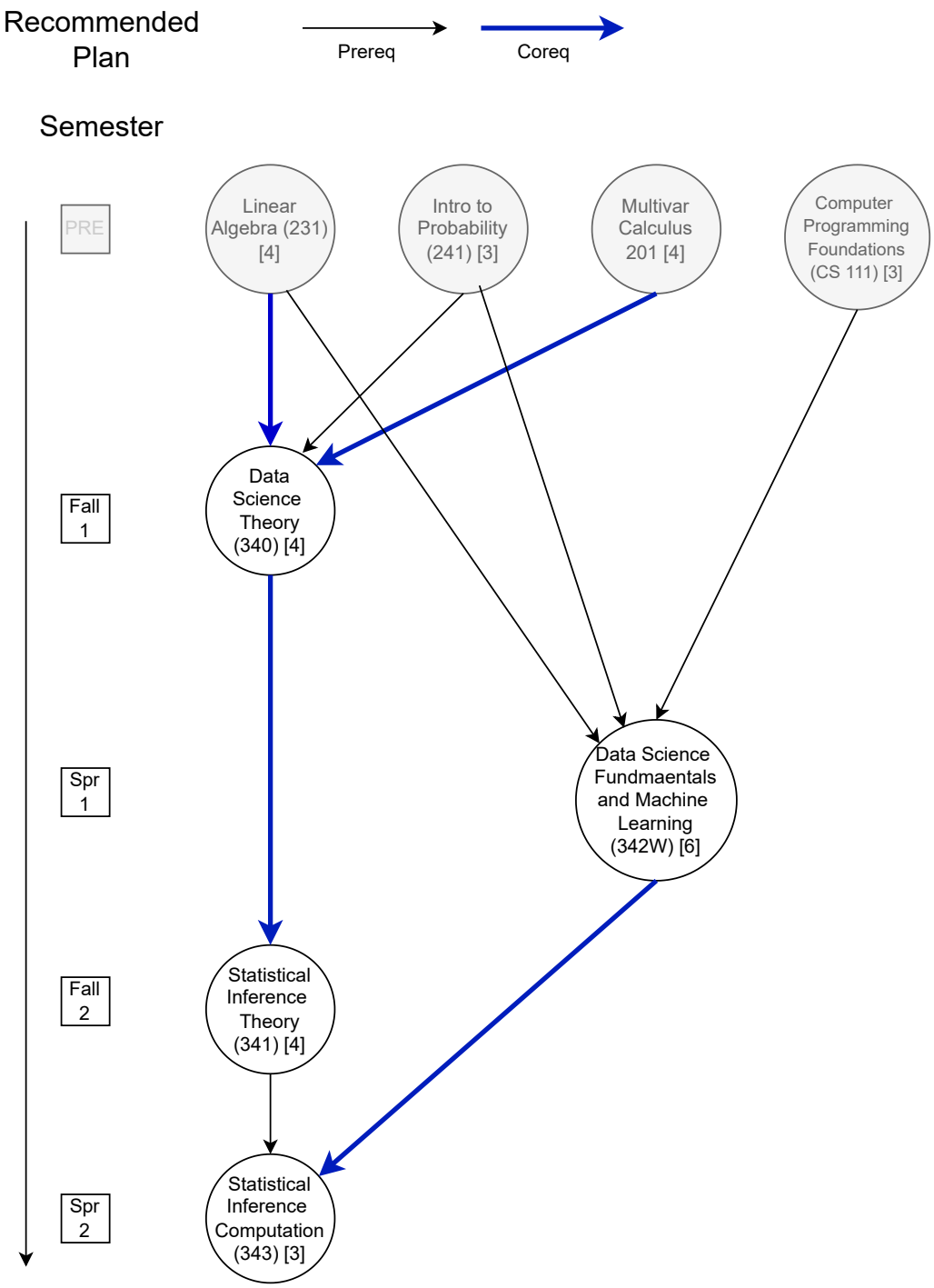
The Four Data Science Core Classes

This course is one of the four data science core courses but it does not cover any typical "data science" topics. Instead, it is designed to provide theoretical skills for Math 341 and 343. Math 340 and 342W are designed to be standalone courses and the other two courses rely on topics covered therein. Thus there is an order the classes need to be taken. Below

are two plans, the first is over two semesters and hence it is not *not recommended* as it will be a *very* heavy workload. The second is over four semesters and it is the recommended plan as I believe it will allow students to absorb the material more effectively:



Examining the above, we note that MATH 341 and 343 form a series of two statistics courses: the first, theoretical with traditional topics and the second, computational with modern topics. Both heavily rely on the theoretical topics taught in MATH 340.


Recommended Plan
Prereq
Coreq
Semester
PRE
Fall 1
Spr 1
Fall 2
Spr 2
Linear Algebra (231) [4]
Intro to Probability (241) [3]
Multivar Calculus 201 [4]
Computer Programming Foundations (CS 111) [3]
Data Science Theory (340) [4]
Data Science Fundmaentals and Machine Learning (342W) [6]
Statistical Inference Theory (341) [4]
Statistical Inference Computation (343) [3]

This is not your typical mathematics course. This course will do lots of modeling of real-world situations using data via the `R` statistical language.

Prerequisites

MATH 341 / 641 or a foundations of Frequentist and Bayesian statistics course. Critical is coverage of hypothesis testing, confidence intervals, credible regions, maximum likelihood. Implicitly, MATH 340 / 640 is also a prerequisite. At times, we will be drawing on these concepts as well so keep those notes handy.

Corequisites

MATH 342W / 642 or a foundations of data science course. Critical is coverage of supervised learning, basic computing, linear regression and logistic regression.

Course Materials

Textbook: I will be referencing Larry Wasserman's *All of Statistics: A concise course in statistical inference* which can be purchased on Amazon and Casella and Berger's *Statistical Inference* which can be purchased on Amazon. There is no excuse not to have these books. They are *required*. However, I will not ususally be teaching "from the book" — most of the material in the class comes from the lecture notes. The textbooks are a way to get "another take" on the material and they will only cover about only half of the material done in class. For the other half, you will have to make use of other resources. I also recommend Rice's *Mathematical Statistics and Data Analysis*, 3rd edition which can be purchased on Amazon as well but I will not reference it during class.

Computer Software: During lectures, there will be demos using `R` which is a free, open source statistical programming language and console. You can download it from: `http://cran.mirrors.hoobly.com/`.As this course is coreq'd with MATH 342W, this course has a lot of programming in `R` for the homeworks.

Calculator: You can use a TI-84, 85, 89 or any calculator which you wish. I strongly suggest you use Wolfram Alpha and its smartphone app.

The Use of Slack and Github as a Learning Management System

As the course homepage is updated (e.g. a new homework assignment is posted), you will hear about it in slack. Each assignment will have its own channel. You can feel free to discuss things with your fellow students there. If you are asking me a question, you must do so in the `#discussions` channel for a general questions or the assignment-specific channel (e.g. `#HW03`) so other students can see the question and benefit from the answer. If you pm me, I will not answer and just ask you to move it to the public channel. Do not be afraid to ask questions. There are many people who will have your same question!

Slack is a wildly successful company that recently got bought by Salesforce because businesses *really* use it. Pretend you are working at one of these businesses: no posting about random stuff; keep things professional!

We will not be using any features of github for learning management. Do not open "issues" on github!

Announcements

Course announcements will be made via slack in the `#general` channel (not on email). I am known to send a few slack messages per week on important issues. The Slack workspace `#general` channel is also connected to the course homepage via an integration. So every time I change the homepage (e.g. to release a new homework or upload notes or a video), you will get a notification.

I can't stress the following enough: if you are not on slack, you will miss all class announcements!!! Slack notifies you when there are messages. You may wish to mute all channels except for #general. See this article for how to do that.

Class Meetings

There are 28 scheduled meetings. Of these, 23 will be lectures, 2 will be midterm exams which are in class and 3 will be review periods during the meeting before the exams (see lecture schedule section above). The exam schedule is given on page 7. The last class of the semester will be rescheduled to be a review period that is conveniently before the final. We will discuss later in the semester.

Homework

There will be 6–9 theory homework assignments and 6–9 practice homework assignments (labs). Homeworks will be assigned and placed on the course homepage and will usually be due a week later in class. Homework will be **graded** out of 100 with extra credit getting scores possibly > 100. I will be doing the grading and will grade an *arbitrary subset of the assignment* which is determined after the homework is handed in.

Homework must be handed in by emailing it to me as a PDF. You must do one of two things:

- **Print out the homework and handwrite your answers in the alotted space for each question. Then scan your homework as a PDF. There are a ton of good photo-PDF apps for both iPhone and droid.**
- **Open the PDF on your device and use a PDF-editing program to electronically handwrite your answers and save the PDF.**

I will NOT accept homework that is not atop the original rendered homework PDF file. Remember to write your name. There are no regrades during this pandemic semester. Homework must be at maximum 5MB.

You are encouraged to seek help from me if you have questions. After class and office hours are good times. You are highly recommended to work with each other and help each other. You must, however, submit your own solutions, *with your own write-up* and in *your own words*. There can be no collaboration on the actual *writing*. Failure to comply will result in severe penalties. The university honor code is something I take seriously and I send people to the Dean every semester for violations.

Philosophy of Homework

Homework is the *most* important part of this course.[1] Success in Statistics and Mathematics courses comes from experience in working with and thinking about the concepts. It's kind of like weightlifting; you have to lift weights to build muscles. My job as an instructor is to provide assistance through your zone of proximal development. With me, you can grow more than you can alone. To this effect, homework problems are color coded green for easy, yellow for harder, red for challenging and purple for extra credit. You need to know how to do all the greens by yourself. If you've been to class and took notes, they are a joke. Yellows and reds: feel free to work with others. Only do extra credits if you have already finished the assignment. The "[Optional]" problems are for extra practice — highly recommended for exam study.

Time Spent on Homework

This is a three credit course. Thus, the amount of work outside of the 2.5hr in-class time per week is 6-9 hours. I will aim for 7.5hr of homework per week on average. You can think of doing the homework well as "sufficient" as *my* job is to ensure that by *you* doing the homework you will study and understand the concepts in the lectures and you won't have all that much to do when the exams roll around.

Late Assignment Policy

Late homework will be penalized 10 points per business day (Monday–Friday save holidays) for a maximum of five days. *Do not ask for extensions*; just hand in the homework late. After five days, **you can hand it in whenever you want** until *the last scheduled class meeting according to the official academic calendar*. As far as I know, this is one of the most lenient and flexible homework policies in college. I realize things come up. Do not abuse this policy; you will fall far, far behind.

LaTeX Homework Bonus Points

Part of good mathematics is its beautiful presentation. Thus, there will be a 1–5 point bonus added to your theory homework grade for typing up your homework using the LaTeX typesetting system based on the elegance of your presentation. The bonus is arbitrarily determined by me.

[1] In one student's observation, I give a "mind-blowing homework" every week.

I recommend using overleaf to write up your homeworks (make sure you upload both the hw#.tex and the preamble.tex file). This has the advantage of (a) not having to install anything on your computer and thus not having to maintain your LaTeX installation (b) allowing easy collaboration with others (c) alway having a backup of your work since it's always on the cloud. If you insist to have LaTeX running on your computer, you can download it for Windows here and for MAC here. For editing and producing PDF's, I recommend TeXworks which can be downloaded here. Please use the LaTeX code provided on the course homepage for each homework assignment.

If you are handing in homework this way, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. The easiest way to use overleaf is to copy the raw text from hwxx.tex and preamble.tex into two new overleaf tex files with the same name. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the "\vspace" command and draw them in after printing or attach them stapled.

Since this is extra credit, do not ask me for help in setting up your computer with LaTeX in class or in office hours. Also, **never share your LaTeX code with other students** — it is cheating and if you are found I will take it seriously.

Homework Extra Credit

There will be many extra credit questions sprinkled throughout the homeworks. They will be worth a variable number of points arbitrarily assigned based on my perceived difficulty of the exercise. Homework scores in the 140's are not unheard of. They are a good boost to your grade; I once had a student go from a B to and A- based on these bonuses.

Examinations

Examinations are solely based on homeworks (which are rooted in the lectures)! If you can do all the green and yellow problems on the homeworks, the exams should not present any challenge. I will *never* give you exam problems on concepts which you have not seen at home on one of the weekly homework assignments. There will be three exams and the schedule is below.

Exam Schedule

- Midterm examination I will be on [see course homepage] with the first review session on the class meeting prior
- Midterm examination II will be on [see course homepage] with a review on the class meeting prior.
- The final examination will be on [see course homepage] with a review TBA.

Exam Policies and Materials

I allow you to bring any calculator you wish but it cannot be your phone. The only other items allowed are pencil and eraser. I do not recommend using pen but it is allowed. Food is not allowed during exams but beverages are allowed.

I also allow "cheat sheets" on examinations. For midterm I, you are allowed to bring *one* 8.5" × 11" sheet of paper (front and back). Two sheets single sided are not allowed. Midterm II, you are allowed to bring *two* cheat sheets. On these sheets of paper you can write anything you would like which you believe will help you on the exam. For the final, you are allowed to bring *three* 8.5" × 11" sheet of paper (front and back). Six sheets single sided are not allowed. I will be handing back the cheat sheets so you can reuse your midterm cheat sheets for the final if you wish.

Cheating on Exams

If I catch you cheating, you can either take a zero on the exam, or you can roll the dice before the University Honor Council who may choose to suspend you.

Missing Exams

There are no make-up exams. If you miss the exam, you get a zero. If you are sick, I need documentation of your visit to a hospital or doctor. Expect me to call the doctor or hospital to verify the legitimacy of your note.

Accommodations for Students with Disabilities

Candidates with disabilities needing academic accommodation should: 1) register with and provide documentation to the Special Services Office, Frese Hall, Room 111; 2) bring a letter indicating the need for accommodation and what type. This should be done during the first week of class. For more information about services available to Queens College candidates, contact: Special Service Office; Director, Miriam Detres-Hickey, Frese Hall, Room 111; 718-997-5870 (Monday – Thursday 8:00 a.m. to 5:00 p.m. & Friday 8:00 a.m. to 4 p.m.).

Class Participation

This portion of your grade is assessed based on your level of interaction during the course lectures e.g. asking / answering questions. Participation on slack counts towards this total.

Grading and Grading Policy

Your course grade will be calculated based on the percentages as follows:

Theory Homework	10%
Labs	10%
Midterm Examination I	20%
Midterm Examination II	20%
Final Examination	35%
Class participation	5%

The semester is split into three periods :

(a) From the beginning until midterm I. Midterm I covers material during this time.

(b) From midterm I to midterm II. Midterm II covers material in this period only.

(c) From midterm II until the final. The final is cumulative and covers all course material.

Each of the periods is assessed evenly. Thus, each period must count the same towards your grade. Since there is 75% of the grade allotted to exams, there is 25% allotted to each period. Thus, the final is upweighted towards the material covered in the third period. In summary, the final will have 5/35 points $\approx$ 14% for the first period's material, 5/35 points $\approx$ 14% for the second period's material and 25/35 points $\approx$ 71% for the last period's material. A good strategy for the final is to just study the material after Midterm II and minimal studying for the previous material.

The Grade Distribution

As this is a small and advanced class, the class curve will be quite generous. I'm expecting approximately 40% A's and 40% B's. If you do your homework and demonstrate understanding on the exams, you should expect to be rewarded with an A or a B. C's are for those who "dropped out" somewhere mid-semester or who cannot demonstrate basic understanding.

Checking your grade and class standing

You can always check your grades in real-time using `https://qc.gradesly.com`. You will enter in your QC ID number (or CUNYfirst email address). I will provide you with your password by email the first week of class.

Auditing

Auditors are welcome. They are encouraged to do all homework assignments. I will even grade them. Note that the university does not allow auditors to take examinations.